Limits to doping in oxides

J. Robertson[1,*] and S. J. Clark[2]
[1]*Engineering Department, Cambridge University, Cambridge, United Kingdom*
[2]*Physics Department, Durham University, United Kingdom*
(Received 19 August 2010; revised manuscript received 10 January 2011; published 28 February 2011)

The chemical trends of limits to doping of many semiconducting metal oxides is analyzed in terms of the formation energies needed to form the compensating defects. The *n*-type oxides are found to have high electron affinities and charge neutrality levels that lie in midgap or the upper part of their gap, whereas *p*-type oxides have small photoionization potentials and charge neutrality levels lying in the lower gap. The doping-limit energy range is found to vary with the bulk free energy of the compound.

DOI: 10.1103/PhysRevB.83.075205 PACS number(s): 71.55.Gs, 61.72.jd, 61.72.uj, 79.60.Bm

I. INTRODUCTION

Metal oxides display a wide range of useful electronic functionality, such as magnetoresistive, superconducting, or multiferroic properties.[1–4] The behavior of free carriers in $SrTiO_3$ induced by polar surface doping has created widespread interest.[2,3] At the same time, there is considerable effort to develop metal oxides such as ZnO as wide-band-gap light-emitting and high-mobility semiconductors,[5,6] to use amorphous semiconducting oxides such as $InGaZnO_x$ as a higher-mobility channel material instead of amorphous silicon in thin-film transistors for large-area electronics,[7,8] and to use the unusual properties of TiO_2 in photoelectrochemistry for energy conversion and environmental cleanup.[9–11] The latter applications require an ability to dope the oxide in a unipolar fashion and preferably in a bipolar fashion. The standard transparent conducting oxides can be heavily doped *n*-type, even in their amorphous phases, because of their ionic bonding,[8,12] in contrast to covalently bonded a-Si, whose doping efficiency is severely impaired in the amorphous phase.[13] Amorphous phases are particularly useful for large-area, low-cost electronics.[8] On the other hand, *p*-type doping of oxides is rare, and there has been a significant effort to find practical *p*-type oxides.[14–19] It is therefore useful to have guidance on the limits to doping of the various oxides in terms of their energy band alignments by using the pinning energy concepts introduced previously by Walukiewicz,[20,21] Zhang *et al.*,[22,23] and Zunger[24] for the tetrahedrally bonded semiconductors.

Here we analyze the systematics of doping in oxide systems using the concept of pinning energy rule, and we derive their *n*-type pinning energy and *p*-type pinning energy. We find that the dopability of oxides can be viewed in terms of the valence band maximum and the conduction band minimum on an absolute energy scale or with respect to a common alignment energy such as the charge neutrality level derived from their band structures. The well-known *n*-type oxides (Zn, In_2O_3, and SnO_2) are found to have conduction band edges deep below the vacuum level, while the known *p*-type oxides (NiO, Cu_2O, $CuAlO_2$, $CuGaO_2$, and $CuCrO_2$) have high-lying valence band edges. The analysis also predicts that TiO_2 and $SrTiO_3$ will be difficult to dope *p*-type.

Three factors can limit the ability to dope a material effectively: a lack of dopant solubility, the dopant level being too deep and unionizable, or the dopant being compensated by native defects. The key limitation is compensation by native defects. This occurs if moving the Fermi energy to a band edge causes the spontaneous formation of compensating defects because the formation energy of that defect has fallen to zero at that Fermi energy. This requires us to know the formation energies of the native defects of the various oxides.

II. METHOD

The defect-formation energies of the various oxides were calculated by the *ab initio* plane wave pseudopotential method. Generally, this would use the local density formalism (LDF) to approximate the electronic exchange-correlation energy. However, this leads to an underestimate of the band gap, which is a particularly severe problem for the transparent conducting oxides. To overcome this problem, we use hybrid density functionals to represent the exchange-correlation energy.[25–29] These are accurate enough to give good band gaps while being efficient enough to be useful for calculations on supercells of up to 100 atoms. They have the advantage that they are true functionals, so that they can be used in energy minimizations.

We have used the screened exchange (sX) hybrid density functional.[25,26] Screened exchange includes a short-range component of Hartree-Fock exchange. The sX functional replaces all the LDA exchange with a Thomas-Fermi screened Coulombic exchange potential, but retains the LDA version of correlation potential,[26]

$$V_{sX}(r,r') = -\sum_i \frac{\psi_i(r)e^{-k_s|r-r'|}\psi_j^*(r)}{|r-r'|} + \varepsilon_{\rm loc}^{LDA}(\rho) - \int V_X^{LDA}(\rho)F(\rho)\rho(r)dr, \quad (1)$$

where, i and j label the electronic bands, k_s is the inverse Thomas-Fermi screening length, $\varepsilon_{\rm loc}$ is the exchange-correlation energy per electron as given by the LDA, and V_x is the nonlocal exchange-correlation energy per electron evaluated in a homogeneous electron gas of density ρ. k_s is evaluated from the valence electron density, as given in Ref. 26, omitting shallow d core electrons. This has recently been implemented for a plane wave basis set in the CASTEP code.[30]

Clark and Robertson[26] have calculated the band gaps of a wide range of semiconductors and insulators by the sX method,

and it is found to give the minimum band gaps for the relevant oxides within 0.1 eV of their experimental values.

To extend the range of oxides covered, we also used defect-formation energy data from other groups who used the closely related Heyd, Scuseria, Ernzerhof (HSE) hybrid functional.[27,28] HSE includes a fraction of the short-ranged component of the Hartree-Fock exchange, with a different screening length to sX. To date, the band gaps found by sX and HSE06 functionals for the various oxides are very similar, both being close to experiment, despite their different functional form. This is partly because the screening length in HSE06 was fitted to optimize the band gaps. Komsa *et al.*[29] have discussed how the band-gap and defect-formation energies in practice vary relatively little with the functional.

This allows the defect-formation energies to be calculated.[31–43] The sX method was used for defect-formation energies of ZnO,[31] HfO_2,[32] TiO_2,[33] and Cu_2O. Defect-formation energies from HSE are used for ZnO by Oba *et al.*,[34] Agoston *et al.*,[35] and Clark *et al.*,[31] for TiO_2 by Janotti *et al.*[36] and Morgan and Watson,[37] for Cu_2O by Scanlon and Watson,[38] and for $CuAlO_2$ by Scanlon and Watson.[40]

The formation energy H_q of the defect of charge q as a function of the Fermi energy (ΔE_F) from the valence band edge E_V and the relative chemical potential ($\Delta\mu$) of element α is expressed as[44]

$$\Delta H_q(\mu, E_F) = E_q - E_H + q(E_V + \Delta E_F) + \sum_\alpha n_\alpha(\mu_{\alpha,0} + \Delta\mu_\alpha), \quad (2)$$

where $\mu_{\alpha,0}$ is reference chemical potential of element α and n_α is the number of atoms of element α, ΔE_F is the Fermi energy with respect to the valence band edge. The charge state and cell size corrections to the defect-formation energies are handled by the method of Lany and Zunger.[44]

The doping-limit energies are later displayed on a band alignment diagram. For this, the band edges of each oxide (a,b) were aligned in either the Schottky or the Bardeen limit. In the Schottky limit, there is no charge transfer across the interface, and the conduction band offset ϕ_n is given by the difference in electron affinities of a and b,

$$\phi_n = (\chi_a - \chi_b), \quad (3)$$

as in the electron affinity rule. More usually, there is charge transfer at the interface. One model, which assumes that the interfaces bond on their nonpolar faces and neglects specific atomic models of the interfaces, proposes that the lineup is controlled by the lineup of charge neutrality levels,[45]

$$\phi_n = (\chi_a - \Phi_{S,a}) - (\chi_b - \Phi_{S,b}) + S(\Phi_{S,a} - \Phi_{S,b}) \quad (4)$$

where S is the Schottky barrier pinning factor, and $S = 0$ corresponds to the strongly pinned or Bardeen limit. χ_a is the electron affinity of oxide a, and $\Phi_{S,a}$ is the charge neutrality level of oxide a. This approximation holds for nonpolar interfaces.

III. RESULTS AND DISCUSSION

Figure 1 plots the formation energy of the most stable charge state of the most stable donor-type and acceptor-type native defects for various oxides as a function of the Fermi energy E_F, for both O-rich and metal-rich (O-poor) conditions. The slope of the lines is the stable charge state at that value of E_F. O-rich conditions correspond to the chemical potential (μ_O) of O in the O_2 molecule, $\mu_O = 0$. The O-poor conditions correspond to μ_O equal to the metal/metal oxide equilibrium, or μ_O equaling the free energy per O of the bulk oxide.

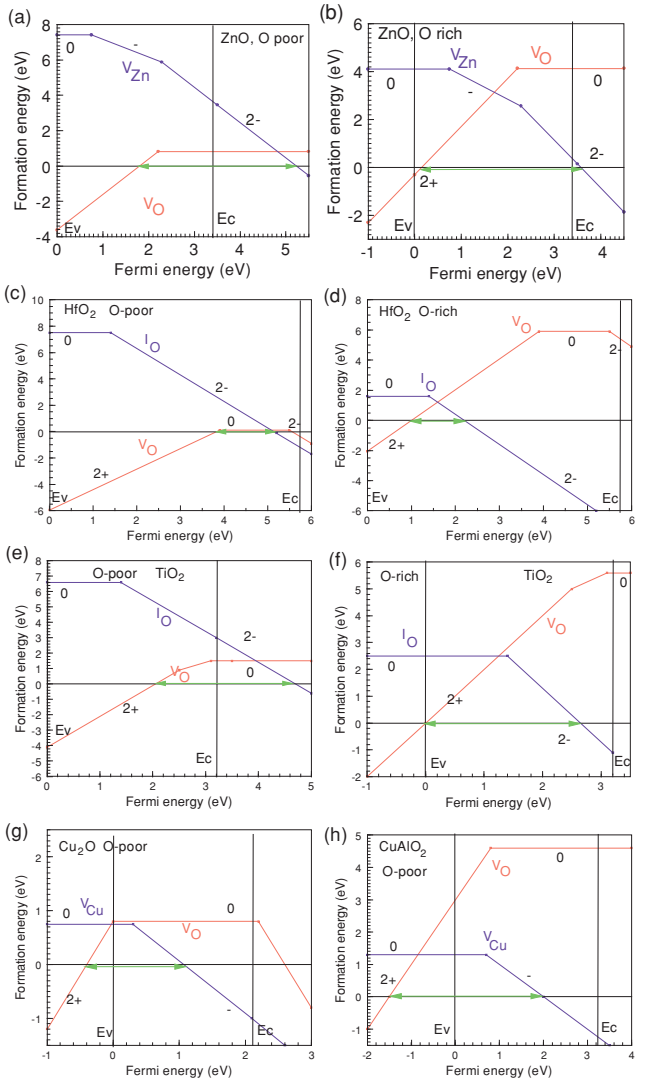

FIG. 1. (Color online) Formation energy of the most stable donor and acceptor native defect vs Fermi energy for (a),(b) an n-type oxide ZnO after Clark *et al.* (Ref. 31), (c),(d) a wide-gap undopable oxide HfO_2 after Xiong *et al.* (Ref. 32), (e),(f) an n-type oxide TiO_2 with data from Janotti *et al.* (Ref. 36), and Morgan *et al.* (Ref. 37), (g) a p-type oxide Cu_2O with data from Scanlon *et al.* (Ref. 38), and p-type $CuAlO_2$ with data from Scanlon *et al.* (Ref. 40).

We apply Eq. (2) to consider the native defects in the case of ZnO, in response to doping. Figure 1(a) shows the formation energies of defects in ZnO by the sX method.[31] (Very similar values are found using HSE.[34]) If a donor (e.g., Al_{Zn}) is used to raise E_F toward the conduction band edge E_C, we must consider the formation energy of possible negatively charged compensating acceptors such as Zn vacancies (V_{Zn}) or oxygen interstitials. Consider first the case of O-poor ZnO, with an O chemical potential $\mu_O = -3.37$ eV. Figure 1(a) shows that V_{Zn} is the more stable of these two defects, but V_{Zn}^{2-} still has

a large positive formation energy (<3 eV) when E_F is at E_C. E_F would have to be raised to 5.2 eV, 1.8 eV above E_C, for Zn vacancies to form spontaneously. This energy is called the "*n*-type pinning energy" for donors.[22]

On the other hand, if an acceptor dopant is used to lower E_F toward the valence band edge E_V, we consider the formation energy of compensating donor defects such as O vacancies V_O or Zn interstitials. V_O is slightly the more stable of the two. Figure 1(a) shows that in the O-poor limit V_O has a negative formation energy if E_F drops below 1.8 eV above E_V. Oxygen vacancies will spontaneously form if E_F moves toward E_V and opposes the *p*-type doping. The 1.8 eV is called the "*p*-type pinning energy." The energy range where we can shift E_F without spontaneously creating compensating defects is between these two limit energies, and equals 3.4 eV.

The same analysis applies to O-rich conditions. Figure 1(b) shows that the *n*- and *p*-type pinning energies both fall by 1.7 eV (this is the calculated free energy of bulk ZnO divided by 2, the charge of the defects). However, the energy range for which any native compensating defect does not form spontaneously is still 3.4 eV, from 0.2 to 3.6 eV. This is one reason that ZnO can only be doped *n*-type. As such, ZnO is representative of a number of conducting oxides such as SnO_2 and In_2O_3.

Figure 2(a) shows the free energy of various bulk oxides per O atom plotted against the work function (electronegativity) of the parent metal.[45] Now consider HfO_2, an oxide that is difficult to dope. It is a highly stable wide band-gap insulator used in electronics.[46] Figure 1(c) shows that its most stable native defects are the O vacancy and the O interstitial. The data in Fig. 1(c) include the sX results of Xiong *et al.*[32] for the O vacancy, and the generalised gradient approximation (GGA) results of Zheng *et al.*[41] for the O interstitial. (Note that the O interstitial in HfO_2 involves only occupied O valence states, which should be adequately handled by GGA.) In O-poor conditions, if we shift E_F toward E_C by *n*-type doping, then the O interstitial I^{2-} will spontaneously form for E_F lying above 5.1 eV, the *n*-type limit, where its formation energy goes negative. On the other hand, if we lower E_F toward E_V, then O vacancies spontaneously form for E_F lying below 3.8 eV, the *p*-type pinning limit. The energy range over which no defects spontaneously form is now quite narrow, only 1.3 eV, from 3.8 to 5.1 eV. In the O-rich case in Fig. 1(d), the pinning limits shift to lower in the band gap, but its magnitude stays the same. HfO_2 illustrates the case of a difficult-to-dope oxide, with a narrow doping range, impossibly far from the band edges.

Figures 1(e) and 1(f) show the case of TiO_2, an important catalytic and electrochemically active oxide with a high bulk free energy. This uses data from the HSE06 results of Janotti *et al.*,[36] the LDA+U results of Morgan and Watson,[37] and some sX results of Clark *et al.*[33] TiO_2 can also represent the energetics of doping in $SrTiO_3$ in that this depends mainly on the TiO_2 sublattice energies, TiO_2 having a lower free energy per O than SrO [Fig. 2(a)]. Figure 1(e) shows the defect-formation energies[36,37] for the O-poor limit at the Ti_2O_3/TiO_2 equilibrium, $\mu_O = -4.07$ eV. In O-poor conditions, if we shift E_F toward E_C, then the possible compensating acceptors, O interstitials, still have a high formation energy. Thus, TiO_2 is easily doped *n*-type. If we lower E_F toward E_V, the formation

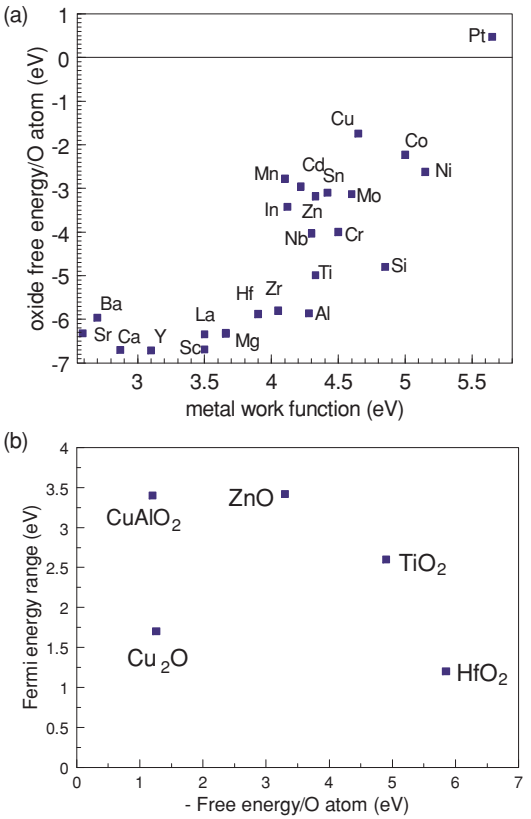

FIG. 2. (Color online) (a) Bulk free energy of oxides per O atom versus the work function of the parent metal atom with data from Robertson *et al.* (Ref. 46). (b) Energy range between *n*-type and *p*-type doping limits vs bulk free energy/O atom, for the various oxides shown in Fig. 1.

energy of donors such as the O vacancy becomes negative for E_F below 2.1 eV above E_V, Fig. 1(e). The energy range between the *n*- and *p*-type pinning limits is 2.7 eV. In O-rich conditions, the pinning limits shift to lower energies, so that the *p*-type pinning limit falls to 0.3 eV above E_V. The O-poor case is more relevant. However, as the *p*-type limit lies above the valence band edge in both the O-poor and the O-rich cases, Figs. 1(e) and 1(f) indicate that it will be very difficult to dope TiO_2 or $SrTiO_3$ *p*-type, without causing compensation, because the *p*-type pinning limit lies so high in its band gap. It is therefore unlikely that $SrTiO_3$ can be made a bipolar semiconductor by electronic doping.

Figure 1(g) shows the case of Cu_2O, a less strongly bonded *p*-type oxide, using the HSE06 data of Scanlon and Watson[38] and the GGA data of Raebinger *et al.*[39] (Raebinger *et al.*[39] applied correction methods to account for the band-gap error; Clark *et al.*[33] finds similar vacancy formation energies for Cu_2O by the sX method.) We see that shifting E_F above 1.1 eV causes the formation energy of Cu vacancy acceptors to become negative. These states compensate any external donors. On the other hand, E_F can be shifted down to E_V by external acceptors without causing any native donor (O vacancy) to have a negative formation energy. Thus, they are uncompensated. The energy range between doping limits is 1.7 eV in Cu_2O. Thus, Fig. 1(g) shows that Cu_2O is possible to dope *p*-type.

Figure 1(h) shows the equivalent doping limits for $CuAlO_2$ using the defect-formation energies calculated by Scanlon and Watson[40] by HSE. They find a band gap of 3.5 eV. The compensating defects are still the O vacancy and the Cu vacancy. The Cu vacancy compensates donors and has similar characteristics to that in Cu_2O, with its formation energy passing through zero at $E_F = 2.0$ eV. On the other hand, the O vacancy is much more stable, with a formation energy of +4.6 eV in its neutral state, compared to only 0.8 eV in Cu_2O. Thus, the O vacancy formation energy only passes below zero for E_F below −1.5 eV. This gives a doping energy range of 3.5 eV, similar to that in ZnO. The critical feature is that the strong Al-O bonds have formed a framework which anchors the O atoms into it. This makes O vacancies more costly and makes doping more difficult to compensate.

Previous analyses of doping limits have implicitly assumed that the doping energy range is approximately constant. Figure 2(b) shows the pinning energy range ΔE vs the oxide free energy per O atom of Fig. 2(a). We see that ΔE is not constant; it follows a chemical trend. ΔE first increases with the bulk free energy of the oxide. This is to be expected as the cost of an oxygen vacancy would increase for oxides with a higher free energy. Then, for the ionic oxides like TiO_2 and HfO_2, ΔE decreases, as O interstitial defects become the lower-cost donor than metal vacancies, and their presence curtails the possibility of *n*-type doping in the oxide. On the other hand, it is possible to obtain a more constant ΔE in a framework structure such as $CuAlO_2$ in which all O's are bonded to some high-cohesive-energy component, while the active component (Cu-O) has a lower cohesive energy.

The pinning-limit energies for the various oxides in the O-poor limit can be assembled into a band diagram, first for bands aligned according at the Schottky limit. This is shown in Fig. 3. We see that the *n*-pinning limit lies at a reasonably constant energy for ZnO, SnO_2, TiO_2, and $CuAlO_2$. Similarly, the *p*-limit energy also lies at a reasonably constant energy for these compounds. (Preliminary calculations suggest that the defect energetics of SnO_2 are similar to those of ZnO and TiO_2.) For Cu_2O and HfO_2 the pinning energies lie closer together.

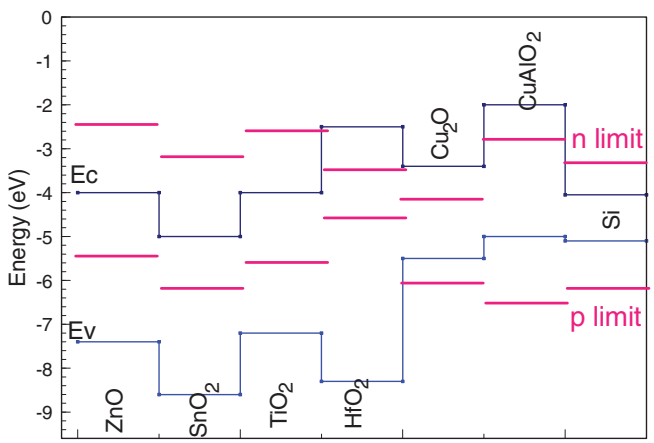

FIG. 3. (a) Doping limits for the O-poor case for the oxides considered in Fig. 1, with their valence and conduction band energies plotted against the vacuum level.

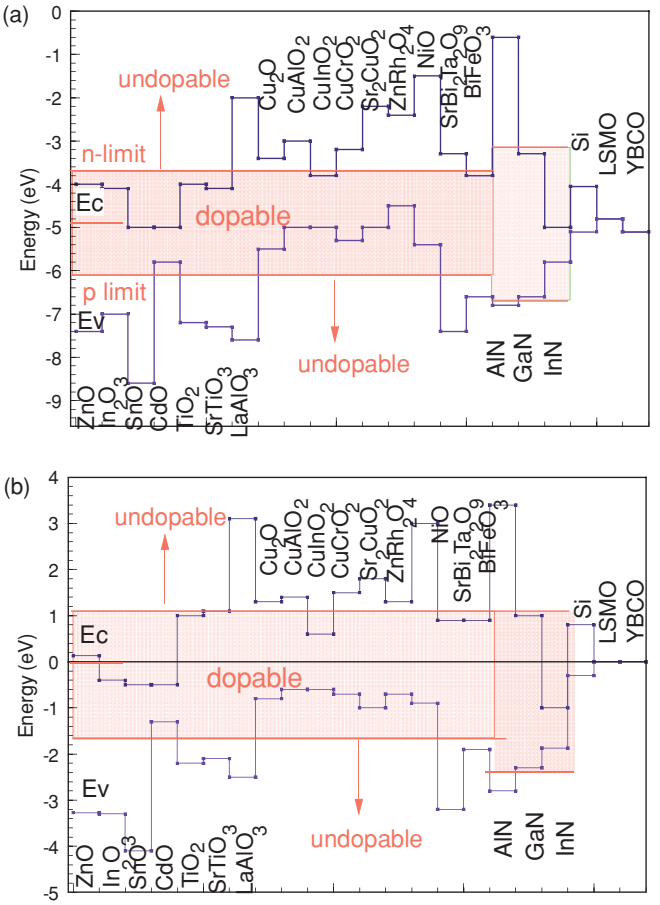

FIG. 4. (a) Valence and conduction band energies of various oxides vs vacuum level, with the doping limits, showing the dopable and undopable cases. (b) Similar plot, with the oxide bands aligned using their charge neutrality levels (CNLs).

It is useful to extend this to a wide range of oxides, to make predictions for oxides for which there are no explicit calculations of defect-formation energies. To do this, we assume that the pinning energies lie at constant energies, as in Walukiewicz,[21] taking the values given in ZnO and $CuAlO_2$ for their O-poor regime. They are plotted on global band diagrams in Fig. 4. Figure 4(a) shows the case where the band offsets of the oxides are shown in the "Schottky limit," using electron affinities and ionization potentials as determined by photoemission or electrochemistry (Table I).[14,47–53] The data for the conducting oxides ZnO, In_2O_3, SnO_2, Cu_2O, $CuAlO_2$, $CuInO_2$, and $SrCu_2O_2$ are from Hosono,[48] those for the nitrides are from Ref. 50, and those for TiO_2 are from Gratzel.[11] The experimental data for $SrTiO_3$ and $LaAlO_3$ are from Chambers *et al.*[51] and Edge *et al.*[52] The $SrTiO_3$:$LaAlO_3$ interface has most of its 2.4-eV-wide band-gap mismatch taken up as a near 2-eV conduction band offset, with a small offset in the valence band, according to Chambers *et al.*[53] The data for $CuCrO_2$ is from Benko.[14] Its band gap is similar to that of $CuAlO_2$.[54] $BiFeO_3$ data is estimated.[55] The work function of $(La,Sr)MnO_3$ (LSMO) is from Ref. 56. Figure 4(b) shows band alignments in the "Bardeen limit," in which screening causes the offsets to align their charge neutrality levels (CNLs). The CNL is the branch point of the complex band structure, where

TABLE I. Electron affinity (eV), minimum band gap (eV), charge neutrality level above valence band edge (eV), and relevant reference for each compound.

	EA	E(gap)	CNL	Ref.
Si	4.0	1.1	0.3	
ZnO	4.0	3.4	3.28	48
In_2O_3	4.1	2.9	3.2	48,62
SnO_2	5.0	3.6	4.1	48
CdO	5.0	0.8	1.3	48
TiO_2	4.0	3.2	2.6	11,45
$SrTiO_3$	4.1	3.2	2.6	45,51
$LaAlO_3$	2.0	5.6	3.1	52,59
Cu_2O	3.4	2.12	0.8	48
$CuAlO_2$	2	3.0	0.8*	48,71
$CuGaO_2$	3	2.1	0.6*	48
$CuInO_2$	3.6	1.4	0.7*	48
$CuCrO_2$	3.0	2.8	0.7*	48,54,72
$SrCu_2O_2$	2.2	2.8	1*	48
Zn_2RhO_4	2.4	2.2	0.7*	48
NiO	1.5	4.3	0.9*	48
$SrBi_2Ta_2O_9$	3.3	4.1	3.2	45
$BiFeO_3$	3.8*	2.8	1.9	55
HfO_2	2.5	5.8	3.7	47
AlN	0.6	6.2	2.8	50,60
GaN	3.7	3.3	2.3	60,64
InN	5	0.8	1.9	50
$LaSrMnO_3$	4.8	0	0	56
YBCO	5.1	0	0	

the character of the gap states changes from valence band to conduction bandlike.[57] Calculated and experimental CNL values are used.[58–63] The chemical trends of band alignments are seen to be similar in both cases.

N-type and *p*-type pinning-limit energies are plotted in Fig. 4. It is convenient at first to treat the pinning limits as independent of the oxide. The explanation of Fig. 4 is as follows. The pinning-limit energies set the range over which E_F can be varied, without causing the formation of compensating defects that stop doping. If the *n*-type limit lies above E_C of that oxide, E_F can be shifted to E_C without defect formation and it can be doped *n*-type. If the *p*-type limit is below E_V, then that oxide can be doped *p*-type without defect formation. Figure 4(a) says that the criterion for *n*-type dopability is that the electron affinity should be large (E_C well below the vacuum level), while *p*-type dopability requires the valence band ionization potential to be small (E_V not too far below the vacuum level). Figure 4(b) says that the criterion is that the CNL should lie in midgap or the upper gap for *n*-type dupability, or at midgap or in the lower gap for *p*-type dopability.

The *n*-type transparent conducting oxides stand out in having large electron affinities and also having CNLs that lie in their upper gap or indeed in the conduction bands. However, the corollary is that ZnO, In_2O_3, and SnO_2 have too-deep valence band edges to be doped *p*-type. TiO_2 and $SrTiO_3$ also fall into this category, their valence band edges are too deep below the vacuum level. The alloying of La with Sr in LSMO to achieve hole doping can be considered to be an interstitial doping of a framework structure. We see that the Fermi level lies within its doping limits so that we do not expect the hole doping to cause a compensation response in this compound.

For *n*-type oxides, these results largely confirm what is already known for ZnO from more detailed studies. *P*-type doping will be highly compensated and is favored by a high-O chemical potential. A way to avoid compensation is to avoid the defect concentrations reaching equilibrium. Also, as noted in Ref. 66, defects and dopants at grain boundaries have different formation energies, which might be exploited to obtain a *p*-doping response.

The overall model shows that the situation can be generalized to similar oxides such as SnO_2, TiO_2, and $SrTiO_3$ even if there is not the same level of detailed calculations for these cases.

The *p*-type oxides Cu_2O, $CuAlO_2$, $CuCrO_2$, $SrCu_2O_2$, and NiO all have low ionization potentials so that their valence band edge lies above the *p*-type limit in Fig. 4(a). This correlates with their *p*-like character. $CuInO_2$ can be doped in both directions because its smaller gap allows its E_C to lie below the *n*-type pinning limit. It shows that the principle of Kawazoe[15] to make *p*-type oxides, of using the Cu d : O $2p$ hybridization to broaden the valence band and lower its effective mass, also works because it raises E_V toward the vacuum level. The prime example of this effect is in NiO caused by Ni $3d$: O $2p$ hybridization. It also occurs in $CuCrO_2$. The defect properties of the *p*-type oxides have only recently been studied. The benefit of the present work is that it applies a global model and shows that they behave in a manner expected from other systems.

The comparison of the greater doping ability of $CuAlO_2$ than that of Cu_2O is valuable. It shows that the principle advantage of $CuAlO_2$ is not just its wider band gap, making it transparent, but that its strongly bonded framework of Al-O bonds gives it much improved defect properties which inhibit dopant compensation. The same basic mechanism was proposed for amorphous $InGaZnO_x$ semiconductors as used in thin-film transistors (TFTs).[8] Ga is added to $InZnO_x$ to increase the cost of O deficiency, to thereby reduce the background conductivity of the undoped material and the off-current of the TFT.

A further question is if wide-gap oxides like ZnO usually only show unipolar doping, why can nitrides (GaN) with the same band gap show bipolar doping?[64] Figure 5 shows the formation energy diagram for the lowest energy defect for GaN. In Ga-rich GaN, the least costly defect turns out to be the N vacancy[65] over the full range of E_F around the gap. This is due to its charge states. The lower slope for V^+ in the formation energy diagram in Fig. 5 leads to a larger ΔE range for GaN than for ZnO. Luckily, this range actually spans the band gap of GaN, rather than being asymmetrically disposed, so that the *n*- and *p*-type pinning limits fall just outside its band gap. It can therefore be doped both *n*- and *p*-type without compensation.

It is also useful to consider the second factor that limits doping, the deepness of the dopant level within the band gap.[66–68] It has recently been noted that the substitutional N site is at least 1 eV deep in ZnO and consistent with ESR data,[69] which showed a localized hole state.[66,67] On the other hand, substitutional N_O is well known to form a deep level in TiO_2.[70] Similarly, substitutional Mg in GaN is only a

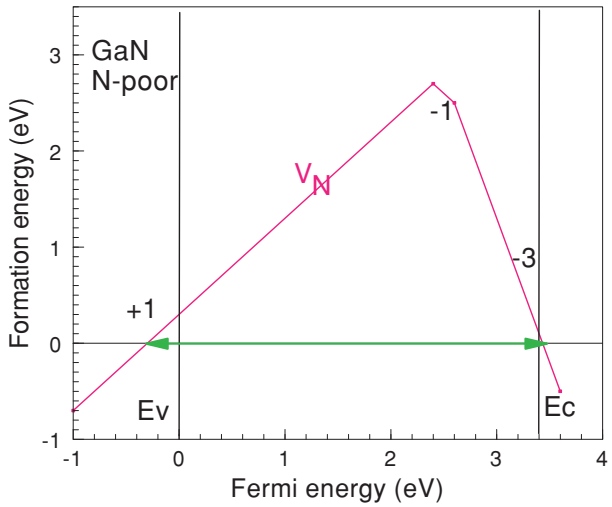

FIG. 5. (Color online) Defect-formation energies for N-poor GaN, showing that its doping limits lie just outside its band gap. Data from Ref. 65.

borderline shallow acceptor which shows metastable state.[67] These cases show that systems which are difficult to dope in terms of compensation can also be difficult to dope in terms of the depth of the levels. Both factors follow the same basic chemical trends.

IV. SUMMARY

The limits to doping in metal oxides have been calculated from their defect-formation energies. The chemical trends of limits to doping of many metal oxides are plotted on a band lineup diagram, in terms of doping pinning energies. Oxides which can be doped *n*-type are found to have high electron affinities and charge neutrality levels that lie in midgap or the upper part of their gap. Oxides that can be doped *p*-type have small photoionization potentials and charge neutrality levels lying in the lower gap. TiO_2 and $SrTiO_3$ have low probability of being able to be doped *p*-type. There are advantages in using complex oxides such as $CuAlO_2$ to inhibit the formation of compensating defects.

ACKNOWLEDGMENTS

The authors thank H. Hosono for drawing attention to Ref. 48.

*jr@eng.cam.ac.uk

[1] C. H. Ahn, J. M. Triscone, and J. Mannhart, Nature (London) **424**, 1015 (2003).
[2] A. Ohtomo and H. Y. Hwang, Nature (London) **427**, 423 (2004).
[3] A. Brinkman, M. Huijben, M. van Zalk, J. Hujiben, U. Zeitler, J. C. Maan, W. G. van der Wiel, G. Rijnders, D. H. A. Blank, and H. Hilgenkamp, Nat. Mater. **6**, 493 (2007).
[4] J. Wang, J. B. Neaton, H. Zheng, V. Nagarajan, S. B. Ogale, B. Liu, D. Viehland, V. Valthyanathan, D. G. Schlom, U. V. Waghmare, N. A. Spaldin, K. M. Rabe, M. Wuttig, and R. Ramesh, Science **299**, 1719 (2003).
[5] D. C. Look, B. Chaflin, Y. I. Allvov, and S. J. Park, Phys. Stat. Solidi. A **201**, 2203 (2004).
[6] A. Tsukazaki, A. Ohtomo, T. Kita, Y. Ohno, and M. Kawasaki, Science **315**, 1388 (2007).
[7] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano, and H. Hosono, Nature (London) **432**, 488 (2004).
[8] H. Hosono, J. Non. Cryst. Solids **352**, 851 (2006); K. Nomura, A. Takagi, T. Kamiya, and H. Ohta, Jpn. J. Appl. Phys. **45**, 4303 (2006).
[9] A. Fujishima, X. Zhang, and D. A. Tryk, Surf. Sci Rep. **63**, 515 (2008).
[10] U. Diebold, Surf. Sci. Rep. **48**, 53 (2003).
[11] M. Gratzel, Nature (London) **414**, 338 (2001).
[12] J. Robertson, Phys. Stat. Solidi B **245**, 1026 (2009).
[13] R. A. Street, Phys. Rev. Lett. **49**, 1187 (1982).
[14] F. A. Benko and F. P. Koffyberg, Mater. Res. Bull. **21**, 753 (1986).
[15] H. Kawazoe, N. Yasukawa, H. Hyodo, M. Kurita, H. Yanagi, and H. Hosono, Nature (London) **389**, 939 (1997).
[16] H. Yanagi, T. Hase, S. Ibuki, K. Ueda, and H. Hosono, Appl. Phys. Lett. **78**, 1583 (2001).
[17] A. Kudo, H. Yanagi, H. Hosono, and H. Kawazoe, Appl. Phys. Lett. **73**, 220 (1998).
[18] H. Ohta, M. Orita, M. Hirano, I. Yagi, K. Ueda, and H. Hosono, J. Appl. Phys. **91**, 3074 (2002).
[19] R. Nagarajan, A. D. Draeseke, A. W. Sleight, and J. Tate, J. Appl. Phys. **89**, 8022 (2001).
[20] W. Walukiewicz, Appl. Phys. Lett. **54**, 2094 (1989); J. Cryst. Growth **159**, 244 (1996).
[21] W. Walukiewicz, Physica B **302-303**, 123 (2001).
[22] S. B. Zhang, S. H. Wei, and A. Zunger, J. Appl. Phys. **83**, 3192 (1998).
[23] S. B. Zhang, S. H. Wei, and A. Zunger, Phys. Rev. Lett. **84**, 1232 (2000); S. B. Zhang, J. Phys.: Condens. Matter **14**, R881 (2002).
[24] A. Zunger, Appl. Phys. Lett. **83**, 57 (2003).
[25] D. M. Bylander and L. Kleinman, Phys. Rev. B **41**, 7868 (1990).
[26] S. J. Clark and J. Robertson, Phys. Rev. B **82**, 085208 (2010).
[27] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).
[28] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. **125**, 224106 (2006).
[29] H. P. Komsa, P. Broqvist, and A. Pasquarello, Phys. Rev. B **81**, 205118 (2010).
[30] S. J. Clark, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, Z. Krist. **220**, 567 (2005).
[31] S. J. Clark, J. Robertson, S. Lany, and A. Zunger, Phys. Rev. B **81**, 115311 (2010).
[32] K. Xiong, J. Robertson, M. Gibson, and S. J. Clark, Appl. Phys. Lett. **87**, 283505 (2005).
[33] S. J. Clark, H. Y. Lee, and J. Robertson (unpublished).
[34] F. Oba, A. Togo, I. Tanaka, J. Paier, and G. Kresse, Phys. Rev. B **77**, 245202 (2008).
[35] P. Agoston, K. Albe, R. M. Nieminen, and M. J. Puska, Phys. Rev. Lett. **103**, 245501 (2009).
[36] A. Janotti, J. B. Varley, P. Rinke, N. Umezawa, G. Kresse, and C. G. van de Walle, Phys. Rev. B **81**, 085212 (2010).

[37]B. J. Morgan and G. W. Watson, Phys. Rev. B **80**, 233102 (2009).
[38]D. O. Scanlon and G. W. Watson, J. Phys. Chem. Lett. **1**, 2582 (2010).
[39]H. Raebinger, S. Lany, and A. Zunger, Phys. Rev. B **76**, 045209 (2007).
[40]D. O. Scanlon and G. W. Watson, J. Phys. Chem. Lett. **1**, 3195 (2010).
[41]J. X. Zheng, G. Ceder, T. Maxisch, W. K. Chim, and W. K. Choi, Phys. Rev. B **75**, 104112 (2007).
[42]R. K. Astala and P. D. Bristowe, Modell. Simul. Mater. Sci. Proc. **12**, 79 (2004).
[43]A. K. Singh, A. Janotti, M. Scheffler, and C. G. van de Walle, Phys. Rev. Lett. **101**, 055502 (2008).
[44]S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).
[45]J. Robertson, J. Vac. Sci. Technol. B **18**, 1785 (2000).
[46]J. Robertson, O. Sharia, and A. A. Demkov, Appl. Phys. Lett. **91**, 132912 (2007).
[47]J. Robertson, Rep. Prog. Phys. **69**, 327 (2006).
[48]H. Hosono, in *Recent Progress in Transparent Electronics*, edited by A. Facchetti and T. Marks (Wiley, New York, 2010), Chap. 2.
[49]W. Schmickler and J. W. Schultze, in *Modern Aspects of Electrochemistry,* edited by J. M. O'Bockris (Plenum, London, 1986), Vol. 17.
[50]S. P. Grabowski, M. Schneider, H. Nienhaus, W. Monch, R. Dimitrov, O. Ambacher, and M. Stutzmann, Appl. Phys. Lett. **78**, 2503 (2001).
[51]S. A. Chambers, Y. Liang, Z. Yu, R. Droopad, J. Ramdani, and K. Eisenbeiser, Appl. Phys. Lett. **77**, 1662 (2000).
[52]L. F. Edge, D. G. Schlom, S. A. Chambers, E. Cicerrella, J. L. Freeouf, B. Hollander, and J. Schubert, Appl. Phys. Lett. **84**, 726 (2004).
[53]S. A. Chambers, M. H. Englehard, V. Shutthanadan, Z. Zhu, T. C. Droubay, T. Feng, H. D. Lee, T. Gustafsson, E. Garfunkel, A. Shah, J. M. Zuo, and Q. M. Ramasse, Surf. Sci. Rep. **65**, 317 (2010).
[54]D. O. Scanlon, K. G. Godinho, B. J. Morgan, and G. W. Watson, J. Chem. Phys. **132**, 024707 (2010).
[55]S. J. Clark and J. Robertson, Appl. Phys. Lett. **90**, 132903 (2007).
[56]M. Minohara, I. Ohkubo, H. Kumigashira, and M. Oshima, Appl. Phys. Lett. **90**, 132123 (2007).
[57]J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[58]C. G. van de Walle and J. Neugebauer, Nature (London) **423**, 626 (2003).
[59]P. W. Peacock and J. Robertson, J. Appl. Phys. **92**, 4712 (2002).
[60]J. Robertson and B. Falabretti, J. Appl. Phys. **100**, 014111 (2006).
[61]X. Nie, S. H. Wei, and S. B. Zhang, Phys. Rev. Lett. **88**, 066405 (2002).
[62]P. D. C. King, T. D. Veal, D. J. Payne, A. Bourlange, R. G. Egdell, and C. F. McConville, Phys. Rev. Lett. **101**, 116808 (2008); Phys. Rev. B **80**, 081201 (2009); **79**, 035203 (2009).
[63]A. Schleife, F. Fuchs, C. Rodl, J. Furthmuller, and F. Bechstedt, Appl. Phys. Lett. **94**, 012104 (2009).
[64]C. G. van de Walle and J. Neugebauer, J. Appl. Phys. **95**, 3851 (2004).
[65]M. G. Ganchenkova and R. M. Nieminen, Phys. Rev. Lett. **96**, 196402 (2006).
[66]J. L. Lyons, A. Janotti, and C. G. van de Walle, Appl. Phys. Lett. **95**, 252105 (2009).
[67]S. Lany and A. Zunger, Phys. Rev. B **81**, 205209 (2010).
[68]S. Lany and A. Zunger, Appl. Phys. Lett. **96**, 142114 (2010).
[69]W. E. Carlos, E. R. Glaser, and D. C. Look, Physica B **308-310**, 976 (2001).
[70]C. DiValentini, G. Pacchioni, and A. Selloni, Phys. Rev. B **70**, 085116 (2004).
[71]J. Pellicer-Pores, A. Segura, A. S. Gililand, A. Munoz, P. Rodriguez-Hernandez, D. Kim, M. S. Lee, and T. Y. Kim, App. Phys. Lett. **88**, 181904 (2006).
[72]A. C. Rastogi, S. H. Lim, and S. B. Desu, J. Appl. Phys. **104**, 023712 (2008).